



Reduced complexity MPEG2 video post-processing for HD display

Virk, Kamran; Li, Huiying; Forchhammer, Søren

Published in:
IEEE International Conference on Multimedia and Expo, 2008

Link to article, DOI:
[10.1109/ICME.2008.4607548](https://doi.org/10.1109/ICME.2008.4607548)

Publication date:
2008

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Virk, K., Li, H., & Forchhammer, S. (2008). Reduced complexity MPEG2 video post-processing for HD display. In *IEEE International Conference on Multimedia and Expo, 2008* IEEE.
<https://doi.org/10.1109/ICME.2008.4607548>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

REDUCED COMPLEXITY MPEG2 VIDEO POST-PROCESSING FOR HD DISPLAY

K. Virk, H. Li**, S. Forchhammer***

*Technical Univ. Denmark, kam_virk@hotmail.com and **Technical Univ. Denmark, {hli,sf}@com.dtu.dk

ABSTRACT

This paper presents MPEG(2) decoder post-processing for High Definition (HD) flat panel displays [1]. The focus is to design efficient post-processing to reduce blocking and ringing artifacts [2-4]. Standard deblocking modules are improved to obtain a significant load reduction through a new DCT based control scheme. Standard deringing modules are enhanced through adaptive thresholding to improve the image quality. The schemes are implemented in a MPEG2 decoder for evaluation. The enhanced deblocking filter results in load reduction with an overall reduction in execution time of 41~46% over the basic implementation. The enhanced deringing combined with the deblocking achieves PSNR improvements on average of 0.5 dB over the basic deblocking and deringing on SDTV and HDTV test sequences. The deblocking and deringing models described in the paper are generic and applicable to a wide variety of common (8x8) DCT-block based real-time video schemes.

Keywords: DCT—discrete cosine transform, SDTV standard definition television, HDTV-high definition television, MPEG post-processing.

1. INTRODUCTION

Flat panel TV displays are the consumer's choice of today. These displays have a higher resolution and being pixel-based they reproduce pixels distinctly. This leads to a risk of either magnifying MPEG(2) coding artifacts or undesired blurring to avoid these. MPEG-4 p. 10/H.264 [5] provides some tools to reduce the artifacts, but both MPEG-2 and H.264 will co-exist for some time to come.

The goal of this research is to design MPEG(2) decoder post-processing mechanisms for HD displays [6]. The MPEG artifacts due to quantization can be differentiated by visual appearance. We focus on blocking and ringing, which when present are quite annoying in the high definition domain. For Standard Definition TV (SDTV) MPEG video input, the up-scaling may enhance the artifacts. For HDTV MPEG video the complexity becomes an important factor. We address both these issues, focusing on not too complex methods, as e.g. [3][4][7]. In [4] deblocking and deringing is considered for (Q)CIF without exploring SD and HD contents. The PSNR performance is close to standard decoding. In [3] fast deblocking is considered for (Q)CIF as in [4]. We take a different approach of using more MPEG stream information (e.g. motion vectors) in the control and apply it to higher resolution video enhancing the methods in [7].

The rest of the paper is organized as follows: Section 2 briefly presents a standard popular post-processing scheme [7]. Section 3 introduces the proposed enhanced architecture of deblocking and ringing filters. Section 4 presents the results obtained by the proposed methods within an MPEG2 implementation.

2. BASIC ARCHITECTURE

The post-processing filter implementations for MPEG (2) described in this paper are based on the MPEG4 part 2 [7],[8] deblocking and deringing filters. MPEG4 deblocking filters are quite adaptive in nature. Hence the basic architecture involves three distinct execution patterns:

- 1) Deblocking; contains modules of horizontal and vertical deblocking filters along (8x8) block boundaries after reconstruction
- 2) Deringing; contains an (8x8) block based deringing filter introduced after deblocking
- 3) Macroblock level quantization parameter (QP); these values are used as control parameter for the deblocking and deringing filters

MPEG4 p2 deblocking operations are based on two modes of operations at certain block boundary; the first mode is called DC offset mode and this is used for very smooth regions in the video frame, while the other mode is called default mode, operating on complex structural image details. The deblocked video frame becomes an input to the deringing filter. The deringing filter operates by determination of a threshold (by QP), which is used to label pixel regions within a block. And finally the adaptive deringing smoothing filter is applied. This is called the basic post-processing.

In our experiments it is observed that the MPEG4 p. 2 post filtering may result in a decrease in PSNR, when applying deringing at SDTV resolution (PAL). Hence the approach of this work is to further enhance these post-processing modules w.r.t. quality as well as complexity.

3. ENHANCED ARCHITECTURE

It is observed from experiments that the basic MPEG4 p2 based deblocking and deringing filters can be further enhanced by adding some additional controlling parameters. The enhanced filters shall also use information from the MPEG encoded stream such as motion vectors, macroblock type etc to control the post-processing.

Fig. 1 shows the enhanced MPEG2 architecture for post-processing modules. Three additional parameters are added as control along with QP; DCT RUNS represents the number of non-zero DCT

coefficients, MV is calculated from the motion vectors for the macroblock, and MB Type shows macroblock types (I, P, B).

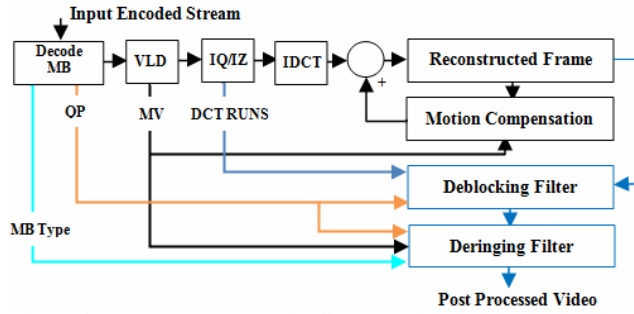


Figure 1: Enhanced composite filter architecture

DCT RUNS will replace the previous MPEG4 based pixel based decision for edge (8 pixels per edge) based mode decision in order to reduce blocking artifacts. MV and MB Type will be used in the deringing filtering in order to improve PSNR values. The post-processing is applied to the luminance component only.

3.1. Enhanced Deblocking

The basic deblocking filter is based on pixel level DC vs. Default mode decisions. This pixel level mode decision leads to expensive CPU loads for HD data contents. To reduce complexity, the enhanced deblocking applies a block level control based on the counts for non-zero discrete cosine transform coefficients (DCT RUNS). This control has replaced the basic DC/(Default) mode decisions control in the reference (basic) implementation.

In the enhanced scheme, the non-zero DCT coefficients are counted for each luminance (8x8) block during decoding. These counts are stored in a DCT RUNS structure as can be seen in Fig. 1. Formally it is defined by

$$dct_runs(k, m_{ii,jj}) = \sum_{n=1}^{64} coeff(k, n) \quad (1)$$

$$coeff(k, n) = \begin{cases} 1 & \text{if } c(k, n) \geq 1 \\ 0 & \text{else} \end{cases}$$

where $k \in \{1, 2, 3, 4\}$ is the block access index of the macroblock, $m_{ii,jj}$ is the macroblock height and width indices, and $c(k, n)$ the DCT coefficient n in k .

These DCT RUNS are utilized for DC/(Default) mode decisions in the horizontal and the vertical deblocking filters. In the horizontal filtering process each pixel at a vertical block boundary is evaluated for DC/(Default) mode. This new DCT RUNS based mode decision is based on block level statistics evaluated at each block boundary (8 pixels per boundary). Each horizontal blocking edge (8 pixels) between two vertical bounding blocks is evaluated for mode decisions by

$$mod e = \begin{cases} DC & \text{if } (dct_runs(k, m_{ii,jj}) < TH \& \& \\ & dct_runs(k+1, m_{ii,jj}) + mb_edge < TH) \\ Default & \text{otherwise} \end{cases} \quad (2)$$

where,

$$mb_edge = \begin{cases} 1 & \text{if } (h_{i,j} \% 16 == 0) \\ 0 & \text{else} \end{cases}$$

where $h_{i,j}$ is the image height and width indices, and $\%$ denotes the modulus operator.

Vertical filtering is defined in the same way, just operating at horizontal block boundaries. Threshold (TH) is currently set to '2' based on experiments for both horizontal and vertical filtering.

It should be noted that it is also necessary to consider the skipped blocks within a macroblock. For Block skip mode we store the DCT RUNS of previously coded blocks. The block level skip mode (CBP) for MPEG2 is also taken care of in the enhanced implementation.

3.2 Enhanced Deringing

The basic deringing implementation (from MPEG4 p2) defines a *max diff* threshold, for labeling the regions, adding a constant value, 4, to the QP. It is known that coded macroblocks have different dynamic characteristics such as light and heavy motion (motion vectors), low and high textures and different coding types (I, P, B). Adding a static value to QP for *max diff* thresholding does not robustly provide good PSNR values. Hence there is a desire to define an adaptive threshold, which will improve the deringing performance.

MPEG4 based deringing uses a constant value 4 to control the *max diff* value, which is obviously not a good measure to exactly estimate the texture information inside a block. Hence a new adaptive threshold scheme is introduced in the deringing filter. To control the *max diff* value, the new scheme introduces two additional parameters along with QP; the macroblock type (MB Type) and the macroblock motion vectors (MV) (Fig. 1). These two parameters will better estimate the threshold for *max diff* according to the macroblock motion vectors and macroblock types, which help the deringing filter decide whether to process or not along with QP values. MB Type (MB_TYPE) represents the motion compensation (I, P, B) information of a macroblock. The MV variable holds the sum of the (average) magnitudes of the predicted motion vectors associated with a certain macro block. Using these values to characterize the dynamic behavior of macroblocks, the adaptive deringing is based on refining the threshold (*max diff*) value by

$$max_diff = \begin{cases} QP & \text{if } (MV / MV_TH \neq 0) \\ CQP & \text{else} \end{cases} \quad (3)$$

where $MV = |PMVX| + |PMVY|$, $PMVX$ and $PMVY$ are the predicted motion vector's x- and y- coordinate, respectively, and $/$ denotes integer division. Motion threshold MV_TH is set as 4 for P frames, 5 for B frames.

$$CQP = \begin{cases} QP & \text{if } (MB_TYPE == INTRA) \\ QP - 1 & \text{else} \end{cases}$$

where *INTRA* is the macroblock type flag and *CQP* is the quantizer value of current macroblock for four (8 x 8) luminance blocks.

4. RESULTS AND DISCUSSION

The deringing and deblocking filters are evaluated both by objective measures and subjective assessment. Tests have been conducted on a series of video sequences including three SD (704 x 576) progressive video sequences ICE, CITY, CREW, and three HD (1280 x 720) sequences Mobcal, Shields, Stockholm. The SD sequences have frame rates of 30f/s. Mobcal, Shields are 50f/s, and the framerate of Stockholm is 60f/s. The MPEG GOP length is *N*=12 and 2 B-frames (*M*=3) between P-frames are used in this test.

4.1 PSNR measurements

The HD sequences are tested at 12M, 15M and 18M bit/s. For SD, CITY, CREW and ICE are tested at 2-6M bit/s. The range of bit-rates is also intended to reflect the variations in bit-rate, which occur in systems where statistical multiplexing is used.

Table 1 shows the overall PSNR gain of the post-processing compared to decoding without post-processing as an average over the range of bit-rates for each of the three SD sequences. The PSNR gain is reported as the difference in PSNR between the results with and without processing. The results show that the MPEG2 enhanced composite filters on average achieve a PSNR gain (E-B) of about 0.5 dB over the composite basic version. The enhanced deblocking (DB) provides a small improvement. The PSNR improvement is mainly due to the enhanced MPEG2 deringing (DR) filter providing a more robust performance than the basic version.

Table 1: Overall PSNR (dB) gain of MPEG2 and MPEG4 post-processing

		ICE	CITY	CREW
DB	Basic	-0.02	-0.06	0.21
	Enhanced	0.22	0.10	0.33
DR	Basic	-0.12	-0.59	-0.07
	Enhanced	0.18	-0.20	0.07
	MoMuSys	0.10	-0.49	0.06
COM	Basic	-0.38	-0.79	-0.07
	Enhanced	0.32	-0.18	0.29
	MoMuSys	0.01	-0.56	0.13
	E - B	0.70	0.61	0.36

Similar results are obtained for HD. The enhanced version achieves PSNR gains of 0.49, 0.47, and 0.45 dB over the basic version for Mobcal, Shields, and Stockholm, respectively.

Table 1 also shows that MPEG4 p2 (MoMuSys) post-filtering [8] is not always robust. On CITY, MoMuSys has a PSNR drop for deringing and composite processing. To analyse the influence of the bit-rate,

Figure 2 depicts the average gain for each of the bit-rates. The figure and the table show very similar relative performance for the post-filtering of MoMuSys MPEG4 video [8] and the performance of these filters when implemented with MPEG2, i.e. the basic version. The deringing (and composite) PSNR performance of the MoMuSys and the basic filters are not robust. The enhanced version improves performance robustly providing a PSNR gain.

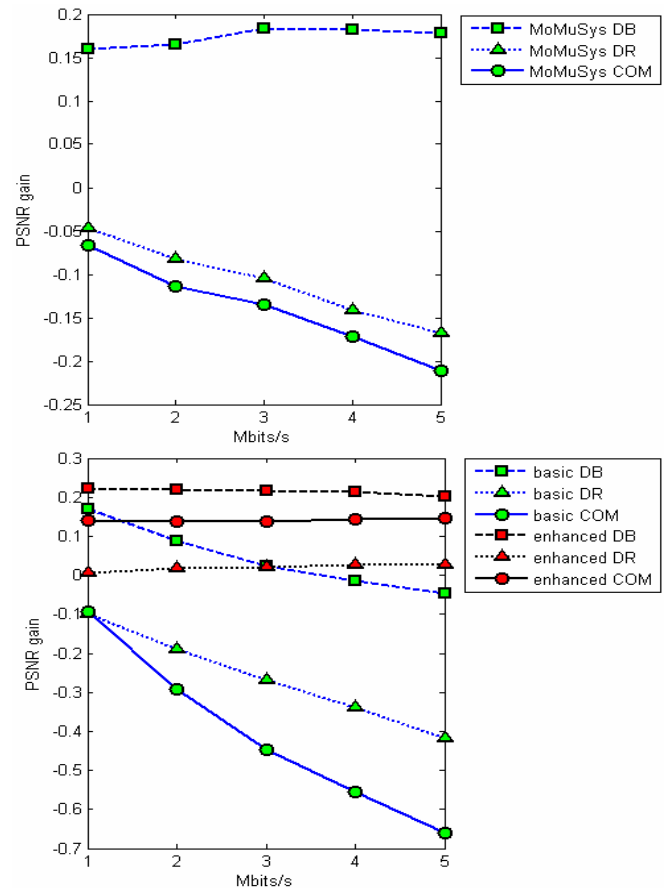


Figure 2: Average PSNR gains for the reference MPEG4 (top) and MPEG2 basic and enhanced filters (bottom) for the CREW, CITY and ICE sequences at different bitrates. PSNR gains are given for deringing (DR), deblocking (DB) and composite (COM) filters.

4.2 Subjective Quality

The post-processed sequences were displayed on a 50" plasma screen for visual assessment. The post-processing reduced the artifacts resulting in a better visual appearance. The enhanced filters introduced very little blurring, thus achieving the best visual results. This is in line with the PSNR performance figures.

Figure 3 shows a comparison of a portion of a frame from the SD sequence FRIES without and with deblocking. The horizontal and vertical blocks appearing in reconstructed image are significantly reduced in deblocked image.



Figure 3: MPEG2 Decoding without (left) and with deblocking (right). Part of frame 38 of FRIES.

Figure 4 shows a comparison of the portion of a frame from the reconstructed and deringing filtered sequence ICE. The ringing artifacts appearing in the left image are significantly reduced. This reduction in artifacts results in a better subjective image.



Figure 4: MPEG2 Decoding without (left) and with deringing (right). Part of frame 18 of ICE.

4.3 Computational Complexity

The deblocking filter is further evaluated from a load complexity point of view. Figure 5 shows a comparison of decision counts for the basic and enhanced deblocking filters. The significant reduction in decision counts for block level mode decisions over the pixel-based decisions are clear for both SD (PAL) and HD resolutions.

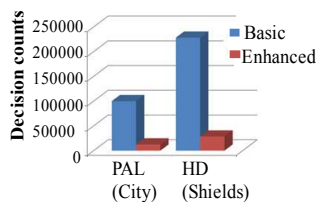


Figure 5: Comparison between pixel based and block based DC and Default mode decisions

Scenes		MPEG2 decoder load reduction over basic
SD	City	42%
	Ice	46%
	Crew	43%
HD	Shields	43%
	Mobcal	41%

Table 2: Load comparison between MPEG2 decoding with basic and enhanced filters

The overall speed on a PC of the MPEG2 basic and enhanced filters are also computed in terms of the full decoding time including post-processing. For this purpose different SD and HD sequences are selected for load complexity comparison. Table 2

presents relative savings for the enhanced MPEG2 post-filtering compared with the basic version over 100 frames. It is clear that the MPEG2 decoder with enhanced filters has achieved a considerable reduction in decoding time reducing the time by 41-46%.

5. CONCLUSION

MPEG(2) video post-processing filters for deblocking and deringing are presented and evaluated. The MPEG4 p2 filters are enhanced by utilizing motion, DCT coefficient and MB type information, besides the quantization parameters. The enhanced filters were implemented in a MPEG2 decoder. The computational performance of the enhanced deblocking filter is improved by replacing pixel level with block level decisions. Adaptive threshold techniques were introduced in the enhanced deringing, improving deringing PSNR and making the deringing more robust. The average PSNR gain of the enhanced composite post-processing is around 0.5 dB compared to the basic filters. Gains in complexity reductions for decoding with enhanced composite filters are quite significant with an overall 41-46% load reduction over decoding using the basic composite filters.

6. REFERENCES

- [1] P. E. Pettersson, "Digital HDTV using MPEG2 Where Are They Now?" *IEE Colloquium on Advanced, Widescreen and High Definition Television Systems*. Issue 6, pp. 4/1 - 4/2, Feb 1996
- [2] Kong, H.-S., Vetro, A., and Sun, H. "Edge map guided adaptive post-filter for blocking and ringing artifacts removal," *IEEE International Symposium on Circuits and Systems (ISCAS)* Vol. 3. 2004
- [3] S.C. Tai, Y.-Y. Chen and S.-F. Sheu, "Deblocking filter for low bit rate MPEG-4 Video," *IEEE Trans. Circ. Syst. Video Tech.* pp. 733-741, June 2005.
- [4] C. Kim, "Adaptive post-filtering for reducing blocking and ringing artifacts in low bit-rate video coding," *Image Communication*, pp. 525-535, 2002.
- [5] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC Video Coding Standard," *IEEE Trans. Circuits Systems Video Tech.* Vol. 13, No. 7, July 2003
- [6] Nakamura, K. Yoshitome, T. and Yashima, Y. "Super high resolution video codec system with multiple MPEG-s HDTV codec LSI's," *Proc. Int'l. Symp. Circuits and Systems 2004*, ISCAS Volume 3, pp. 23-26, 2004.
- [7] MPEG4 Video Standard, Part 2, *Information Technology – Coding of Audio-Visual Objects – Part2: Visual*, International Standards Org. ISO/IEC Int'l Standard 14496-2, 1999.
- [8] MPEG4 Part 2 software, *MoMuSys-FDIS-V1.0-990812* reference software.